



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population

Citation for published version:

Hamilton, DF, Loth, FL, Giesinger, JM, Giesinger, K, MacDonald, D, Patton, JT, Simpson, H & Howie, CR 2017, 'Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population', *The Bone & Joint Journal*, vol. 99-B, no. 2, BJJ-2016-0606, pp. 218-224. <https://doi.org/10.1302/0301-620X.99B2.BJJ-2016-0606.R1>

Digital Object Identifier (DOI):

[10.1302/0301-620X.99B2.BJJ-2016-0606.R1](https://doi.org/10.1302/0301-620X.99B2.BJJ-2016-0606.R1)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

The Bone & Joint Journal

Publisher Rights Statement:

This is the author's final peer-reviewed manuscript as accepted for publication.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



■ ARTHROPLASTY

Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population

D. F. Hamilton,
F. L. Loth,
J. M. Giesinger,
K. Giesinger,
D. J. MacDonald,
J. T. Patton,
A. H. R. W. Simpson,
C. R. Howie

From University of
Edinburgh,
Edinburgh, United
Kingdom

■ D. F. Hamilton, PhD, Research
Fellow, Department of
Orthopaedics and Trauma
■ D. J. MacDonald, BA, MQIP,
Research Co-ordinator,
Department of Orthopaedics and
Trauma

■ J. T. Patton, FRCS Ed,
Consultant Orthopaedic Surgeon,
Department of Orthopaedics and
Trauma
■ C. R. Howie, FRCS Ed, Professor
of Orthopaedics and Trauma,
Department of Orthopaedics and
Trauma
University of Edinburgh,
Edinburgh, UK.

■ F. L. Loth, MSc, PhD candidate,
Department of Psychiatry,
Psychotherapy and
Psychosomatics
■ J. M. Giesinger, PhD, Research
Fellow, Department of Psychiatry,
Psychotherapy and
Psychosomatics
Medical University of Innsbruck,
Innsbruck, Austria.

■ K. Giesinger, MD, MSc,
Consultant Orthopaedic Surgeon,
Department of Orthopaedics and
Traumatology
Kantonsspital St. Gallen, St
Gallen, Switzerland.

■ A. H. R. W. Simpson,
MA(Cantab), DM(Oxon),
FRCS(England & Edinburgh),
George Harrison Law Professor of
Orthopaedic Surgery, Department
of Orthopaedics
Royal Infirmary of Edinburgh, 51
Little France Crescent, Old
Dalkeith Road, Edinburgh, EH16
4SA, UK.

Correspondence should be sent to
D. F. Hamilton; email:
d.f.hamilton@ed.ac.uk

©2017 Hamilton et al
doi:10.1302/0301-620X.99B2.
BJJ-2016-0606.R1 \$2.00

Bone Joint J
2017;99-B:218–24.
Received 22 June 2016; Accepted
after revision 13 October 2016

Aims

To validate the English language Forgotten Joint Score-12 (FJS-12) as a tool to evaluate the outcome of hip and knee arthroplasty in a United Kingdom population.

Patients and Methods

All patients undergoing surgery between January and August 2014 were eligible for inclusion. Prospective data were collected from 205 patients undergoing total hip arthroplasty (THA) and 231 patients undergoing total knee arthroplasty (TKA). Outcomes were assessed with the FJS-12 and the Oxford Hip and Knee Scores (OHS, OKS) pre-operatively, then at six and 12 months post-operatively. Internal consistency, convergent validity, effect size, relative validity and ceiling effects were determined.

Results

Data for the TKA and THA patients showed high internal consistency for the FJS-12 (Cronbach $\alpha = 0.97$ in TKAs, 0.98 in THAs). Convergent validity with the Oxford Scores was high ($r = 0.85$ in TKAs, $r = 0.79$ for THAs). From six to 12 months, the change was higher for the FJS-12 than for the OHS in THA patients (effect size $d = 0.21$ versus -0.03). Ceiling effects at one-year follow-up were low for the FJS-12 with just 3.9% (TKA) and 8.8% (THA) of patients achieving the best possible score.

Conclusion

The FJS-12 has strong measurement properties in terms of validity, internal consistency and sensitivity to change in TKA and THA patients. Low ceiling effects and good relative validity allow the monitoring of longer term outcomes, particularly in well-performing groups after total joint arthroplasty.

Cite this article: *Bone Joint J* 2017;99-B:218–24.

Total joint arthroplasty can achieve high levels of pain relief, function and satisfaction for patients with end-stage arthritis.^{1–3} In the United Kingdom, almost 100 000 hip and knee arthroplasty procedures are undertaken annually.⁴ The success of arthroplasty is assessed by the long-term survival of the implant and by an array of patient-reported outcome measures (PROMs). Traditionally, joint-specific PROMs have evaluated domains such as pain and function. Patient demographics and expectations of ‘reasonable’ post-operative function have changed over the last 20 years: patients now expect to function at ever-higher levels. As such, the ceiling effects of current, commonly used questionnaires may not reflect post-operative changes at this higher level of activity.

Perhaps the ultimate aim of joint replacement surgery is for the patient to ‘forget’ that they have had surgery, or for the outcome to be so good that they become unaware of the (pre-

viously problematic) joint during daily life and activities. The 12 question Forgotten Joint Score (FJS-12)⁵ was introduced in 2012 with the aim of assessing the patients’ ‘joint awareness’. The authors of the score suggest this is more representative of higher-level function after surgery, as to be able to forget about the joint requires the absence of pain and the ability to perform all desired functional tasks without limitation. It is clearly important that an outcome score is psychometrically validated for use and known to accurately and reliably assess the factors that it pertains to. Other groups have used various language translations of the FJS-12 to evaluate outcomes in series of arthroplasty patients.^{6–12} A few authors have made limited attempts at psychometric validation,^{6,8} in particular the reliability of the score to achieve the same result has been evaluated and re-test reliability has been shown to be high, with intraclass correlation

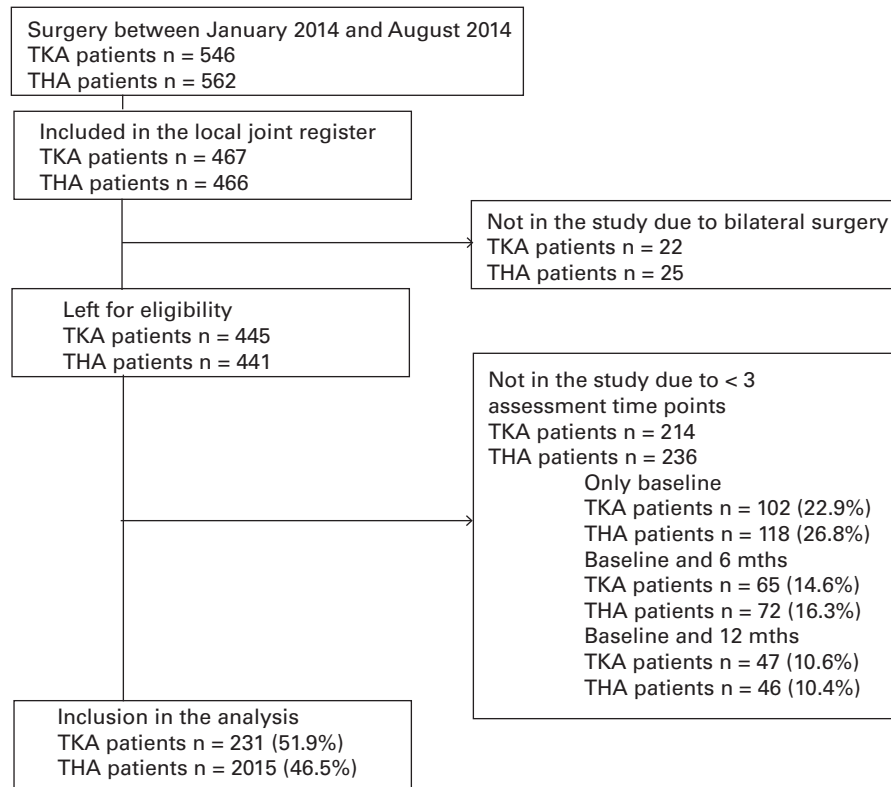


Fig. 1

Patient selection flowchart. TKA, total knee arthroplasty; THA, total hip arthroplasty.

coefficients above 0.9.^{8,13,14} In general, previous studies have only assessed post-operative time points.⁷⁻¹⁷ Only one study, a French translation, has reported the pre-operative data that is essential to infer change from pre-operative score parameters and quantify operative success.⁶ To date, detailed psychometric validation of the score in a native English speaking population is lacking.

The aim of this study was to perform a comprehensive psychometric evaluation and investigate the measurement characteristics (dimensionality, validity and responsiveness) of the English language version of the FJS-12 in total hip and knee arthroplasty patients (THA and TKA). We also sought to validate longitudinal data in contrast to the routinely used Oxford Scores^{18,19} in a United Kingdom population.

Patients and Methods

Participants. Between January 2014 and August 2014, we prospectively assessed 1108 patients who were to undergo unilateral THA or TKA at a single NHS orthopaedic teaching hospital. The study centre is the only hospital to receive adult referrals in a predominantly urban regional population of approximately 850 000 people. Ethical approval was obtained for this project from the institutional review board (11/AL/0079); and patients were recruited with informed consent. All patient undergoing lower limb arthroplasty were eligible for the study. For analysis we required all three assessment time points, thus patients not

providing complete follow-up assessment data were excluded. Individual patients could only be entered into the analysis once to prevent bias, and the second operation in cases of staged bilateral procedures was excluded. There were no other exclusion criteria.

A total of 546 patients underwent TKA and 562 patients underwent THA between January 2014 and August 2014. Prospective data was available for 467 TKA patients and 466 THA patients. A total of 22 TKA patients and 25 THA patients were excluded from the analysis as they underwent bilateral surgery within the study period. As such 445 TKA patients and 441 THA patients were eligible for the study. Of these, 214 TKA patients and 236 THA patients were excluded for not providing questionnaire data at all three assessment time points (pre-operatively, 6-months and 12-months). This resulted in 231 TKA and 205 THA patients eligible for the study (Fig. 1). Excluded patients did not differ significantly with regards to gender (TKA $p = 0.248$ and THA $p = 1.000$) or age (TKA $p = 0.198$ and THA $p = 0.053$).

Demographic data is presented in Table I. The mean age of THA cohort was 69.9 years with a 45.5% male and 55.5% female gender split. The mean age of the TKA cohort was 67.6 years with a 41.5% male and 58.5% female gender split.

Assessments took place at a pre-operative assessment clinic and then by postal questionnaire at six and

Table 1. Patient characteristics at baseline (pre-operatively); total knee arthroplasty n = 231, total hip arthroplasty n = 205

	Age (mean (SD), range)	Gender (M:F)	Side (left:right:missing)
Total knee arthroplasty	69.9 (9.6), 36 to 91	44.2%:55.8% (n = 102:n = 129)	52.4%:40.7%:6.9% (n = 121:n = 94:n = 16)
Total hip arthroplasty	67.6 (11.7), 31 to 92	41.5%:58.5% (n = 85:n = 120)	41.5%:58.5%:0% (n = 85:n = 120:n = 0)

SD, standard deviation

12 months post-operatively. Sociodemographic data was collated prior to surgery. The participants completed outcome questionnaires independently. The FJS-12 questionnaire and the Oxford Knee or Hip Score (OKS or OHS) were completed at all assessment time points. The short-form (SF)-12 questionnaire²⁰ was evaluated 12 months post-operatively to place the joint specific scores in the context of broader health domains.

Assessment instruments. The FJS-12 is a PROM which assesses joint awareness during the activities of daily living (for example, climbing stairs, walking for more than 15 minutes, in bed at night etc).⁵ It consists of 12 questions with a five-point Likert response format.⁵ Item scores are summed and linearly transformed to a 0 to 100 scale, a high value reflecting the ability of the patient to forget about the affected/replaced joint during the activities of daily living.

The OHS and OKS each consist of 12 questions which assess the patient's pain and function.^{18,19} Each item is answered on a five point response scale ranging from 0 to 4, and generates a summed total score ranging from 0 to 48, where 0 indicates the worst possible outcome and 48 good joint function.

The SF-12 Health survey is a 12-item quality of life questionnaire often used in large population health surveys.²⁰ Each item is rated using a five-point Likert response format. The transformed scores for each health domain reflect a patient's poor (0) to excellent (100) health status. Two separate outputs are generated, a physical component summary score (PCS) and mental component summary score (MCS). Normative-based scoring is used for each component: the population mean score is 50 (standard deviation (SD) 10).

Statistical analysis. Sample characteristics are given as means with SD, ranges, and frequencies as appropriate. Dimensionality was explored by means of Cronbach's alpha, exploratory factor analysis (principal component analysis), and confirmatory factor analysis based on polychoric correlations.²¹ Model-data fit for a one-factor solution was assessed using the following statistics and thresholds for adequate fit: root mean square error of approximation (RMSEA)²² < 0.10, Comparative Fit Index (CFI)²² > 0.90, and Tucker-Lewis Index (TLI)²³ > 0.90. Convergent validity was assessed with Pearson correlations between FJS-12 measures and OHS or OKS and the SF-12 Physical Component score. P-values of 0.05 or less were considered as statistically significant. Discriminant validity between the FJS-12 and the SF-12 MCS was assessed using Pearson correlation. Correlation coefficients $r > 0.50$ were

considered as indicator of convergent validity and correlations $r < 0.35$ as an indicator of divergent validity.²⁴

Responsiveness was calculated by the change between pre-operative and six-month follow-up time points, and between the six and 12-month follow-up time points. This is reported as effect size (ES) for the mean change in terms of Cohen's d. In addition, we calculated relative validity (RV) for the FJS-12 and the Oxford Scores. RV (obtained from the ratio of the F-statistics from a repeated measures analysis of variance) gives the ratio of the sample sizes required to detect a mean difference with either of the two PROM instruments. The OKS and OHS Scores were used as a reference measure at all three time points as they have been proven to be responsive in the measurement of change.

Floor and ceiling effects for the FJS-12 and the Oxford Scores are given as the frequency of extreme scores (0 or 100 points for the FJS-12, 0 or 48 points for the Oxford Scores) at the three different time points. In addition, we give the percentage of patients scoring in the extreme 10% of the scale range, similar to the analysis of Jette et al,²⁵ to provide information on the proportion of patients for whom a minimal clinically important change would exceed the range of the scale, and as a consequence would not be measurable. Furthermore, we report the proportion of missing responses on item-level as an indication of response bias. Confirmatory factor analysis was conducted in the software package R (R Foundation for Statistical Computing, Vienna, Austria) using the "lavaan" package.²⁶ All other statistical analyses were performed using SPSS 21.0 (IBM Corp., Armonk, New York).

Results

Dimensionality. Cronbach's alpha coefficients showed high internal consistency of the FJS-12 in both groups ($\alpha = 0.97$ in TKA patients, $\alpha = 0.98$ in THA patients). Exploratory factor analysis of TKA and THA patients suggested a one-factor model for the data, with a single factor explaining 75.6% of variance in TKA patients and 79.6% in THA patients. Eigen values²¹ of further factors were all below 0.68 in both groups. Confirmatory factor analysis showed good model-data fit for a one-factor solution in TKA patients (CFI and TLI > 0.99, RMSEA 0.075) as well as THA patients (CFI and TLI > 0.99, RMSEA 0.084). Standardised factor loadings ranged from 0.82 (item 12) to 0.96 (item 11) in TKA patients and from 0.85 (item 01) to 0.96 (item 11) in THA patients.

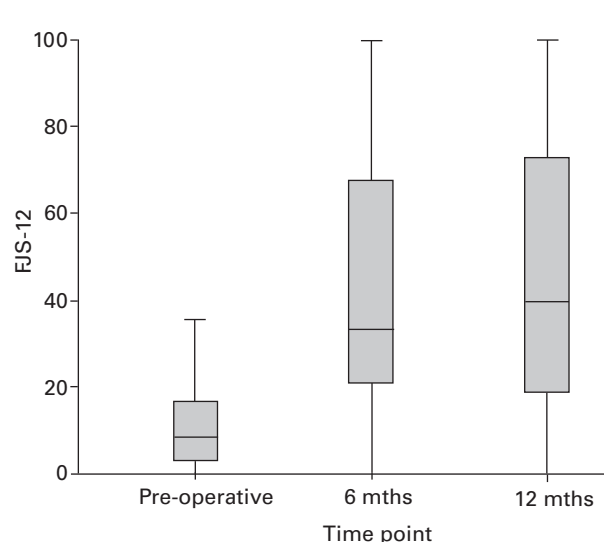


Fig. 2a

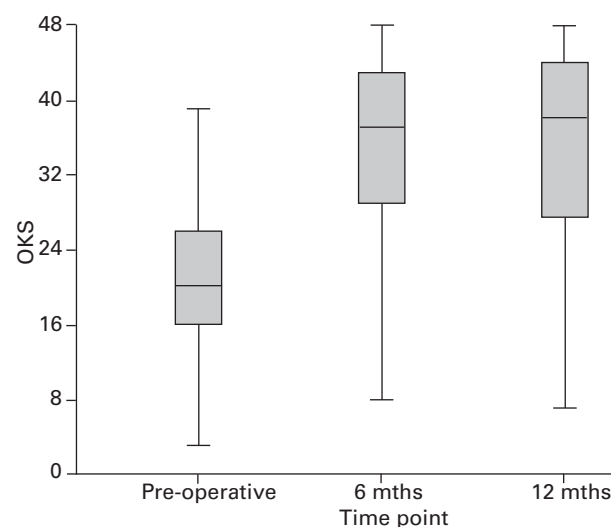


Fig. 2b

Forgotten Joint Score (FJS-12) (a) and Oxford Knee Score (OKS) (b) pre-operatively and six and 12 months post-operatively (knee).

Table II. Means, standard deviations (SD) and responsiveness for the Forgotten Joint Score (FJS-12), Oxford Knee and Hip Scores (OKS and OHS) and the Mental and Physical Component Scores (MCS, PCS)

Pre-operative					6 mths			1 yr			Pre-operative to 6 mths		6 mths to 1 yr	
		Mean	SD	n	Mean	SD	n	Mean	SD	n	ES	RV	ES	RV
Knee	FJS-12	11.5	11.6	231	42.1	30.0	231	45.7	31.2	231	2.65	2.27	0.12	0.31
	OKS	20.7	7.4	231	35.0	8.8	231	35.5	9.3	231	1.95	1	0.06	1
	MCS*	-	-	-	-	-	-	49.0	7.5	230	-	-	-	-
	PCS*	-	-	-	-	-	-	41.2	10.8	230	-	-	-	-
Hip	FJS-12	12.3	15.9	205	48.4	28.7	205	54.3	31.7	205	2.27	2.01	0.21	0.04
	OHS	21.0	9.0	205	38.6	8.6	205	38.3	9.5	205	1.95	1	-0.03	1
	MCS*	-	-	-	-	-	-	47.8	7.6	204	-	-	-	-
	PCS*	-	-	-	-	-	-	43.1	10.8	204	-	-	-	-

* only administered at 12 months
ES, effect size; RV, relative validity

Convergent and discriminant validity. To assess convergent validity we correlated the FJS-12 with the Oxford Scores and the SF-12 PCS. In TKA patients we found high correlations for the OKS ($r = 0.85$) and the SF-12 PCS ($r = 0.70$). Correlations in THA patients were slightly lower; $r = 0.79$ for the OHS and $r = 0.67$ for the SF-12 PCS. Discriminant validity was better in the TKA group (correlation with SF-12 MCS $r = 0.23$) than in the THA group (correlation with SF-12 MCS $r = 0.36$).

Sensitivity to change over time. To measure the performance of a score over time we analysed the data for all time-intervals following surgery (Figs 2 and 3). Results are detailed in Table II.

Pre-operative to six months post-operatively. Results for the FJS-12 showed a large ES for change from pre-operative to six-month follow-up in TKA patients (2.65) and in THA patients (2.27). ES for the OKS and OHS were 1.95 for both TKA patients and THA patients. RV for the FJS-12 was 2.27 in TKA patients and 2.01 in THA patients.

Six months to 12 months post-operatively. ES for the interval between the six month and one-year time question-

naires were smaller for FJS-12 than for the first follow-up (0.12 for TKA patients, 0.21 for THA patients). RV scores decreased to 0.31 in TKA patients and to 0.04 in THA patients. Results for the Oxford Scores showed an ES of 0.06 in TKA patients and -0.03 in THA patients.

Floor and ceiling effects. Pre-operatively, the FJS-12 showed pronounced floor effects, with 15.6% of the TKA patients and 22.4% of the THA patients obtaining the minimum score of 0 points. At one-year follow-up, 3.9% of TKA patients and 8.8% of THA patients achieved the maximum score of 100 points. In contrast, the Oxford Scores did not show floor effects pre-operatively; no patients obtained a score of 0 in either group. At one-year follow-up the maximum score of 48 points was achieved in 0.9% of TKA patients and 14.1% of THA patients.

In a more detailed analysis we defined the floor and ceiling as being within the lowest and highest 10% of the score range.²⁵ This showed pronounced floor effects for the FJS-12 pre-operatively (50.2% of TKA patients and 59.0% of THA patients scored 0 to 10 points). The Oxford Scores were not affected by floor effects. The ceiling effects for the

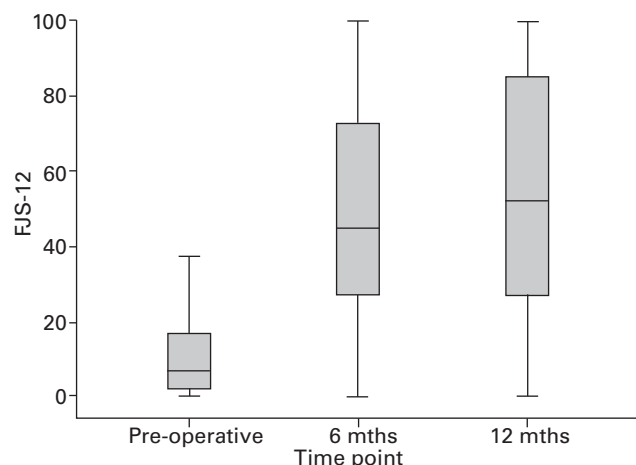


Fig. 3a

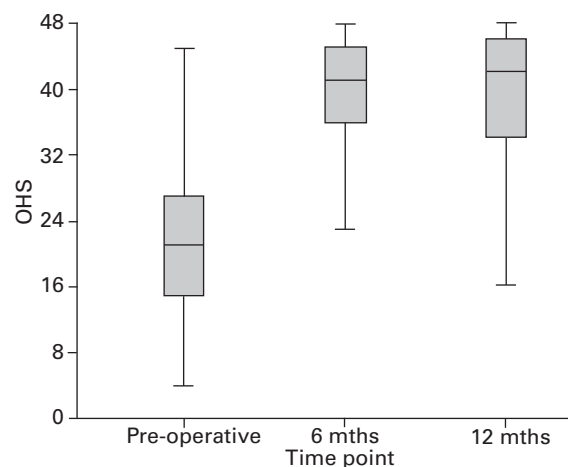


Fig. 3b

Forgotten Joint Score (FJS-12) (a) and Oxford Hip Score (OHS) (b) pre-operatively and six and 12 months post-operatively (hip).

Table III. Floor and ceiling effects of total knee arthroplasty and total hip arthroplasty outcome measures

		Pre-operative (%)				6 mths (%)				1 yr (%)			
		Floor	Floor (10%)	Ceiling	Ceiling (10%)	Floor	Floor (10%)	Ceiling	Ceiling (10%)	Floor	Floor (10%)	Ceiling	Ceiling (10%)
Knee	FJS-12	15.6	50.2	0.0	0.0	3.9	14.3	3.5	9.1	3.0	12.6	3.9	12.6
	OKS	0.0	0.9	0.0	0.0	0.0	0.0	0.9	19.0	0.0	0.0	0.9	25.5
Hip	FJS-12	22.4	59.0	0.0	0.0	2.4	10.2	3.9	9.8	2.0	8.3	8.8	18.5
	OHS	0.0	1.0	0.5	1.0	0.0	1.0	7.3	33.7	0.0	0.0	14.1	39.5

FJS-12, Forgotten Joint Score; OKS, Oxford Knee Score; OHS, Oxford Hip Score

FJS-12 at one year post-operatively were approximately half of those seen in the Oxford Scores at the same time point (THA 18.5% FJS-12 *versus* 39.5% OHS, TKA 12.6% FJS-12 *versus* 25.5% OKS) (Table III).

Missing responses. Analysis of missing responses at a questionnaire item level highlighted generally high levels of completion. Item completion was 98% for the FJS-12 in the TKA cohort with the exception of question 12 (ability to forget about the joint when taking part in sports activities) which had missing responses of 12.1%. The same was found in the THA cohort, with more than 97% item completion for each question except item 12 concerning sporting activity, which was omitted in 19.8% of cases. The Oxford Scores did not suffer from missing responses, with more than 98% item completion across the 12 questions.

Discussion

‘Joint awareness’ is a distinct concept. Qualitative interviews with patients after knee arthroplasty have shown that joint awareness is triggered at a cortical level by pain, functional impairment, and associated sensations (such as numbness, lack of sensation, ‘strange feelings’ with touch, or weather-related sensations).²⁷ Hudak et al²⁸ define this as ‘disrupted embodiment’, where the joint has become separated from the body. From a theoretical standpoint, a high level of post-operative function may be related to the

patient's ability to ‘reunite’ the joint with the rest of the body and to become ‘unaware’ of it in everyday life, akin to the ‘normal’ condition of a ‘healthy’ joint.

With a variety of potential outcome assessment tools available, it is essential that detailed psychometric evaluation is available to inform choice of PROM instruments for the assessment of specific study endpoints. The Consensus-based Standards for the selection of health Measurement Instruments criteria suggest that PROM instruments should meet three main quality domains: reliability, validity and responsiveness.^{29,30}

In this native English speaking population, the FJS-12 shows strong measurement properties in terms of dimensionality (reliability), convergent/discriminant validity and sensitivity to change over time (responsibility).

The Oxford Scores are regularly used tools that have been central to the development of arthroplasty outcomes. They are valid and reliable, although they may suffer from ceiling effects^{31,32} and may therefore lack the sensitivity to capture relevant changes in patient function over longitudinal time periods. As such there has been much interest in developing more sensitive tools with which to assess the outcome of arthroplasty.^{12,33,34}

A major difference between the FJS-12 and the Oxford Scores is the population in which they were designed. The Oxford Scores were developed in pre-operative populations^{18,19}

and reflect the symptom state at that time point. In contrast, the FJS-12 was designed in a post-operative population and the construct reflects the symptom state after joint replacement surgery. This may explain the respective floor and ceiling effects of the scores. We have shown that the Oxford Scores are more responsive in reflecting the initial change between the pre-operative presentation and six months post-operatively, but less so for the change between six- and 12-month follow-up. As such, the FJS-12 had higher effect sizes than the Oxford Scores covering this time period, but lower relative validity. This is directly related to the floor effect of the FJS-12 pre-operatively which results in a low SD of the score at that time point. As the SD is the denominator of Cohen's *d*, the effect sizes are inflated. By contrast, for the later follow-up period (six to 12 months), the FJS-12 outperformed the Oxford Scores in both THA and TKA cohorts. A recent Danish study also found that at later follow-up time points, the OKS has a higher ceiling effect than the FJS-12.¹⁴

To date, this is the most comprehensive evaluation of the psychometric properties of the FJS-12. Others have evaluated selected measurement characteristics in various cohorts and languages. Matsumoto et al⁷ found good convergent validity in a Japanese cohort of hip arthroplasty patients between the FJS-12 and the Western Ontario McMaster Universities Osteoarthritis index (WOMAC) ($r = 0.53$) and a culture-specific hip outcome instrument, the Japanese Orthopaedic Association Hip Disease Evaluation Questionnaire ($r = 0.63$). Thompson et al⁸ found a very high test-retest reliability for the FJS-12 ($r = 0.97$) in an Australian knee arthroplasty cohort. In this study, convergent validity was high with the WOMAC total score ($r = 0.70$) and Knee injury and Osteoarthritis Outcome Score (KOOS) and Quality of Life, Pain and activities of daily living scores ($r = 0.65$). They further reported less ceiling effect in the FJS-12 compared with the WOMAC and KOOS questionnaires. Similarly, Giesinger et al¹¹ indicated superior responsiveness to change for the FJS-12 between one and two year follow-ups compared with the WOMAC and the Knee Society Score, using data from a Swiss clinical trial in knee arthroplasty patients. These three studies broadly support the results of this evaluation, where the FJS-12 showed strong convergent validity with the OKS and OHS and the physical component score of the SF-12. Reassuringly, the association with the mental health component of the SF-12 was low, highlighting appropriate divergent validity.

We report a lesser improvement in the FJS-12 from the pre-operative to one-year post-operative scores compared with Thienpont et al,⁶ who used a French translation of the score in a Belgian population. In their study of 75 THA and 75 TKA patients, there was a more pronounced ceiling effect post-operatively but less of a floor effect pre-operatively. This may highlight a variation in cultural response in international samples, or may also be due to higher rates of missing responses for individual FJS-12 items in that study.

Reassuringly the baseline and post-operative scores in our sample, as measured by the Oxford Scores, are comparable with data from the United Kingdom National Joint Registry,³⁵ suggesting our cohort data to be reflective of the population of the United Kingdom.

In our study, missing responses for individual FJS-12 items were below 3% for 11 out of 12 items in both hip and knee samples. A single question, that of joint awareness during sports, was less well reported. While this suggests that a modest proportion of United Kingdom arthroplasty patients do not consider a question on sporting participation to be applicable, over 80% did respond, perhaps reflecting the evolving expectation of higher level function of patients after joint arthroplasty.

The strengths of this study are its prospective design, a well-defined cohort of patients and the incorporation of pre-operative data. A potential limitation is that this study has a limited follow-up period of 12 months, although this is the typical time period reported in clinical outcome evaluations and in arthroplasty joint registries. The range of comparator outcome metrics is somewhat restricted. A further limitation is the percentage of patients with missing follow-up assessment data at six months or 12 months that were not included in our analysis. While this does not affect the validity of the analyses presented here, it may limit this data in terms of absolute clinical outcome scores.

The current study and previous work suggest that the FJS-12 has the benefits of improved responsiveness and reduced ceiling effects. The low ceiling effects of the FJS-12 may offer advantages when evaluating high performing groups or in powering clinical trials. We suggest that joint awareness could be seen as a relevant outcome domain in its own right, and complementary to questionnaires that assess pain and physical function. It may be that a battery of scores is required to evaluate the full impact of joint surgery from the point of view of the patient. As such, we suggest future studies evaluate qualitative aspects of the patient's conceptualisation of 'joint awareness' and investigate the relative importance of joint awareness compared with joint-specific outcome domains (such as pain and function). Such questions could also be investigated quantitatively using structural equation modelling techniques. It may also be interesting to validate the FJS-12 for other joint conditions to investigate the potential of the score in patient groups where a good outcome following a surgical or conservative intervention is likely to be achieved, for example the conservative management of sports injuries.

The FJS-12 demonstrates strong measurement properties in terms of validity, uni-dimensionality and sensitivity to change in both hip and knee arthroplasty patients. Low ceiling effects and good relative validity suggest that this score may be beneficial for monitoring outcomes over the long-term in patient cohorts expected to achieve high levels of function, or for powering clinical trials.



Take home message:

- This paper validates a new outcome score that assesses the patient's awareness of their joint replacement in a United Kingdom population.

- Strong measurement properties suggest this to be an interesting avenue for further research into patient outcomes, particularly in differentiating outcomes in highly performing patient groups.

Author contributions:

D. F. Hamilton: Study design, Data analysis, Manuscript preparation/editing.
F. L. Loth: Data analysis, Manuscript preparation/editing.
J. M. Giesinger: Study design, Data analysis, Manuscript preparation/editing.
K. Giesinger: Study design, Manuscript editing.
D. J. Macdonald: Study design, Data collection, Manuscript editing.
J. T. Patton: Study design, Manuscript editing.
A. H. R. W. Simpson: Study design, Manuscript editing.
C. R. Howie: Study design, Manuscript editing.

The author or one or more of the authors have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article.

This is an open-access article distributed under the terms of the Creative Commons Attribution licence (CC-BY-NC), which permits unrestricted use, distribution, and reproduction in any medium, but not for commercial gain, provided the original author and source are credited.

This article was primary edited by E. Moulder and first proof edited by A. C. Ross.

References

- Baker PN, van der Meulen JH, Lewsey J, Gregg PJ, National Joint Registry for England and Wales. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. *J Bone Joint Surg [Br]* 2007;89-B:893–900.
- Hamilton DF, Lane JV, Gaston P, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. *BMJ Open* 2013;3:e002525.
- Hamilton DF, Burnett R, Patton JT, et al. Implant design influences patient outcome after total knee arthroplasty: a prospective double-blind randomised controlled trial. *Bone Joint J* 2015;97-B:64–70.
- No authors listed. National Joint Registry. <http://www.njrcentre.org.uk> (date last accessed 19 October 2016).
- Behrend H, Giesinger K, Giesinger JM, Kuster MS. The “forgotten joint” as the ultimate goal in joint arthroplasty: validation of a new patient-reported outcome measure. *J Arthroplasty* 2012;27:430–436.
- Thienpont E, Vanden Berghe A, Schwab PE, Forthomme JP, Cornu O. Joint awareness in osteoarthritis of the hip and knee evaluated with the ‘Forgotten Joint’ Score before and after joint replacement. *Knee Surg Sports Traumatol Arthrosc* 2016;24:3346–3351.
- Matsumoto M, Baba T, Homma Y, et al. Validation study of the Forgotten Joint Score-12 as a universal patient-reported outcome measure. *Eur J Orthop Surg Traumatol* 2015;25:1141–1145.
- Thompson SM, Salmon LJ, Webb JM, Pinczewski LA, Roe JP. Construct Validity and Test Re-Test Reliability of the Forgotten Joint Score. *J Arthroplasty* 2015;30:1902–1905.
- Giesinger K, Hamilton DF, Jost B, Holzner B, Giesinger JM. Comparative responsiveness of outcome measures for total knee arthroplasty. *Osteoarthritis Cartilage* 2014;22:184–189.
- Thienpont E, Zorman D. Higher forgotten joint score for fixed-bearing than for mobile-bearing total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2016;24:2641–2645.
- Thienpont E, Opsomer G, Koninckx A, Houssiau F. Joint awareness in different types of knee arthroplasty evaluated with the Forgotten Joint score. *J Arthroplasty* 2014;29:48–51.
- Bell SW, Stoddard J, Bennett C, London NJ. Accuracy and early outcomes in medial unicompartmental knee arthroplasty performed using patient specific instrumentation. *Knee* 2014;21:S33–S36.
- Shadid MB, Vinken NS, Marting LN, Wolterbeek N. The Dutch version of the Forgotten Joint Score: test-retesting reliability and validation. *Acta Orthop Belg* 2016;82:112–118.
- Thomsen MG, Latifi R, Kallemose T, et al. Good validity and reliability of the forgotten joint score in evaluating the outcome of total knee arthroplasty. *Acta Orthop* 2016;87:280–285.
- Fabre-Aubrespy M, Ollivier M, Pesenti S, Parratte S, Argenson JN. Unicompartmental Knee Arthroplasty in Patients Older Than 75 Results in Better Clinical Outcomes and Similar Survivorship Compared to Total Knee Arthroplasty. A Matched Controlled Study. *J Arthroplasty* 2016;31:2668–2671.
- Thomsen MG, Latifi R, Kallemose T, Husted H, Troelsen A. Does knee awareness differ between different knee arthroplasty prostheses? A matched, case-control, cross-sectional study. *BMC Musculoskelet Disord* 2016;17:141.
- Nielsen KA, Thomsen MG, Latifi R, et al. Does post-operative knee awareness differ between knees in bilateral simultaneous total knee arthroplasty? Predictors of high or low knee awareness. *Knee Surg Sports Traumatol Arthrosc* 2016;24:3352–3358.
- Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg [Br]* 1998;80-B:63–69.
- Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg [Br]* 1996;78-B:185–190.
- Ware J Jr, Kosinski M, Keller SD. A 12-item Short-Form Health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220–233.
- Fayers P, Hays R. *Assessing Quality of Life in Clinical Trials - Methods and practice*. Oxford: Oxford University Press, 2005.
- Browne MW, Cudeck R. Alternative Ways of Assessing Model Fit. *Sociological Methods & Research* 1992;21:230–258.
- Kline RB. *Principles and practice of structural equation modeling*. Second ed. New York: The Guilford Press, 2005.
- Juniper EF, Guyatt GH, Jaeschke R. How to develop and validate a new health-related instrument. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia: Lippincott-Raven Publishers, 1996:49–56.
- Jette AM, Haley SM, Tao W, et al. Prospective evaluation of the AM-PAC-CAT in outpatient rehabilitation settings. *Phys Ther* 2007;87:385–398.
- Rosseel Y. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software* 2012;48:1–36.
- Lewis S, Price M, Dwyer KA, et al. Development of a scale to assess performance following primary total knee arthroplasty. *Value Health* 2014;17:350–359.
- Hudak PL, McKeever PD, Wright JG. Understanding the meaning of satisfaction with treatment outcome. *Med Care* 2004;42:718–725.
- Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–549.
- Mokkink LB, Prinsen CA, Bouter LM, Vet HC, Terwee CB. The COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther* 2016;20:105–113.
- Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651–657.
- Ostendorf M, van Stel HF, Buskens E, et al. Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. *J Bone Joint Surg [Br]* 2004;86-B:801–808.
- Dawson J, Beard DJ, McKibbin H, et al. Development of a patient-reported outcome measure of activity and participation (the OKS-APQ) to supplement the Oxford knee score. *Bone Joint J* 2014;96-B:332–338.
- Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 1998;28:88–96.
- The NJR Editorial Board. 12th Annual Report. National Joint Registry for England, Wales, Northern Ireland and the Isle of Man 2015. <http://www.njrcentre.org.uk/njr-centre/Reports,PublicationsandMinutes/Annualreports/tabid/86/Default.aspx> (date last accessed 19 October 2016).